IEICE Proceeding Series

A Study on Stochastic Animal Swarm Optimization with gradient estimation methods

Takeshi Uchitane, Taro Fukutomi, Toshiharu Hatanaka, Atsushi Yagi

Vol. 2 pp. 306-309
Publication Date: 2014/03/18
Online ISSN: 2188-5079

A Study on Stochastic Animal Swarm Optimization with gradient estimation methods

Takeshi Uchitane[†], Taro Fukutomi[††], Toshiharu Hatanaka[††] and Atsushi Yagi[††]

†Advanced Institute for Computational Science, Riken
Kobe, Hyogo 650-0047, Japan
††Department of Information and Physical Sciences, Osaka University
Suita, Osaka 565-0871, Japan
Email: takeshi.uchitane@riken.jp

Abstract—In this paper, we addressed a search algorithm based on a mathematical swarming model described by stochastic differential equations. The swarm model is constructed by the attractive and repulsive forces as the relationship among search points that imitated animate beings. Numerical simulations are performed to shown availabilities of the proposed method for function optimization.

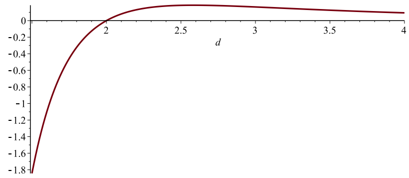

Figure 1: The shape of interaction function h ($r = 2, D = 1$)

1. Introduction

Many population based optimization algorithms whose concepts are based on social animal behaviors have been developed. For example, ant colony optimization employs the concept based on a line of ants [1], artificial bee colony employs the concept based on foraging behavior of honey bees [2] and bacteria foraging optimization employs the concept based on chemotaxis of bacteria [3]. Especially, algorithms whose concepts are based on animal swarming behaviors are called Swarm Intelligence. For representative example, Particle Swarm Optimization developed by Kennedy et. al. in 1995 [4] is one of them.

On the other hand, mathematical swarming models have been studied, e.g. Boids model [5]. In most of such study, behavior of individuals in the swarm is expressed by continuous–time nonlinear dynamical systems and there are analysis and/or numerical study according to such behavior. However there is few study from a viewpoint to develop a optimization algorithms based on a mathematical swarming model described by stochastic differential equations. We believe that the knowledge to develop and/or analyze such model is useful to develop optimization algorithms and the knowledge also is useful to apply such optimization algorithms to real world problems. In order to develop an optimization algorithm by using mathematical swarming model, we start to study about an optimization algorithm based on stochastic fish schooling model [6].

In this paper, we introduce stochastic fish schooling model, then we carry out numerical simulation to show the behavior of fish school (i.e. search points) with an objective function as potential field in order to discuss an ability to search local minimum. Note that our model uses the values of objective function's gradient. Since it is often difficult to know the objective function's shape and/or its gradient values in the real world problems, it is necessary to discuss the the availability to search local minimum when the gradient of potential function is not available. So we also show some model behavior with gradient value estimations based on simultaneous perturbation method [7] and based on function approximation by using radial basis function network [8].

2. Stochastic Fish Schooling Model

We consider that search points are driven by stochastic fish schooling model [6]. In the model, fish has

- no leader,
- common rule to move,
- territory(it's radius is r),
- attraction and repulsion power,
- uncertainty.

First, we consider that $\boldsymbol{x}_i$ is position of i th fish, and the interaction force from j th fish is given by

$$h(\boldsymbol{x}_i, \boldsymbol{x}_j)(\boldsymbol{x}_i - \boldsymbol{x}_j). \tag{1}$$

Here $h(\boldsymbol{x}_i, \boldsymbol{x}_j)$ is a function according to interaction of a pare fish and the summation of interaction force is given by

$$\sum_{j=1, j\neq i}^{N} h(\boldsymbol{x}_i, \boldsymbol{x}_j)(\boldsymbol{x}_i - \boldsymbol{x}_j). \tag{2}$$

Second, we consider that fish move to a direction which the outside potential function $f(\boldsymbol{x})$ is low. Thus, fish has a force $-\gamma\nabla f(\boldsymbol{x}_i)$ which is proportional to the gradient of $f(\boldsymbol{x})$.

Third, we consider that $\sigma_i d\omega_i$ is added as a noise to position of i th fish.

Finally, the velocity of i th fish is denoted by $\boldsymbol{v}_i$, then the dynamics of the fish behavior is given by

$$\frac{d\boldsymbol{x}_i}{dt} = \boldsymbol{v}_i + \sigma_i \frac{d\omega_i}{dt} \quad (3)$$

$$\frac{d\boldsymbol{v}_i}{dt} = -\alpha \sum_{j=1, j\neq i}^{N} h(\boldsymbol{x}_i, \boldsymbol{x}_j)(\boldsymbol{x}_i - \boldsymbol{x}_j) - \gamma \nabla f(\boldsymbol{x}_i). \quad (4)$$

Here $\alpha > 0$ and $\gamma > 0$ are constant.

It is difficult to solve such system by using analytic technique. The feater of such system is often shown by numerical experiments. By using the scheme of Euler method, we get the discrete time model denoted by

$$\boldsymbol{x}_i(t + \Delta t) = \boldsymbol{x}_i(t) + \boldsymbol{v}(t + \Delta t)\Delta t + \sigma \epsilon_i(\Delta t) \quad (5)$$

$$\boldsymbol{v}(t + \Delta t) = \omega \boldsymbol{v}_i(t) - \Big(\alpha \sum_{j=1, j\neq i}^{N} h(\boldsymbol{x}_i(t), \boldsymbol{x}_j(t)) (\boldsymbol{x}_i(t) - \boldsymbol{x}_j(t)) - \gamma \nabla f(\boldsymbol{x}_i(t))\Big)\Delta t. \quad (6)$$

Here ω is inertia constant.

Additionally, we consider gentle force to attract and big force to repulse each other and we have

$$h(\boldsymbol{x}_i, \boldsymbol{x}_j) = \frac{1}{(\|\boldsymbol{x}_i - \boldsymbol{x}_j\|/r)^p} - \frac{1}{(\|\boldsymbol{x}_i - \boldsymbol{x}_j\|/r)^q}. \quad (7)$$

The force $h(\boldsymbol{x}_i, \boldsymbol{x}_j)$ is positive when $\|\boldsymbol{x}_i - \boldsymbol{x}_j\| > r$ and it means that the i th fish and the j th fish are attracted each other. On the other hand the force $h(\boldsymbol{x}_i, \boldsymbol{x}_j)$ is negative when $\|\boldsymbol{x}_i - \boldsymbol{x}_j\| < r$ and it means that the i th fish and the j th fish move away from each other. Moreover, p and q are integer constants with $p > q$. These index numbers make a balance between attractive and repulsive forces.

3. Numerical Simulation of Fish Schooling Model

Here we employ Double Corn function($f_{dc}(\cdot)$) as $f(\boldsymbol{x})$ in equation (6). $f_{dc}(\cdot)$ has two local minimum and it is defined as

$$f_{DC} = \sum_{k=1}^{2} \left(1 - \frac{1}{b_k \|\boldsymbol{x} - \boldsymbol{c}_k\| + 1}\right),$$

where we set $b_1 = 1, b_2 = 2, \boldsymbol{c}_1 = [-2, \ldots, -2]^T, \boldsymbol{c}_2 = [4, \ldots, 4]^T$. Fig.2 shows 2–dimensional contour map.

In order to understand the dynamics of the model, we show the behavior of search points with several values of α and r. The time step size is set as $\Delta t = 0.1$ and final step size is $T = 4000$. In the initial step, the positions of search points are randomly sampled from uniform distribution [-10, 10] (an example of the positions is shown in Fig.2), initial velocity is set as 0, the other parameters are set as $N = 10$, $(p, q) = (3, 5)$, $\sigma = 0.01$, $\omega = 0.9$ and $\gamma = 5$.

Fig.3 and fig.4 show the positions of the fish in the final step with $D = 2$, $\alpha = 1, 2$, $r = 1, 2$. In each case,

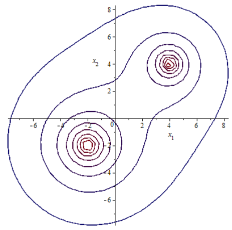

Figure 2: 2–dimensional contour map of Double Cone function

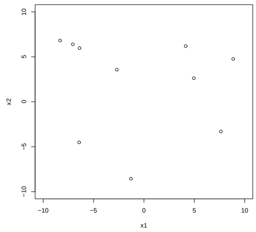

Figure 3: Initial position with $D = 2$.

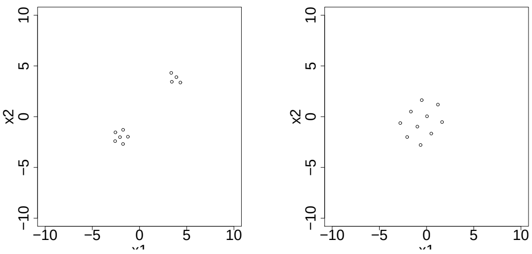

Figure 4: Positions in the final step with $D = 2$, $\alpha = 1$. $r = 1$(left), $r = 2$(right).

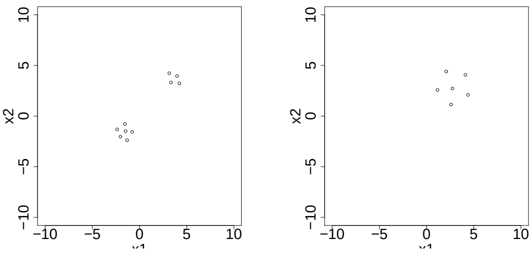

Figure 5: Positions in the final step with $D = 2$, $\alpha = 2$. $r = 1$(left), $r = 2$(right).

search points gather to the near by local minimums. Then we find some cases that population is divided into two local minimums or population gathers to ether local minimum. It seems that population gathers to ether local minimum when r is larger, since the movement of population become stable with large distances between each search points. Which to gather is decided by the balance between interaction force and gradient effects.

4. Applying to function with unavailable gradient value

Generally, population based optimization algorithm is required to find better solutions without gradient of objective function but functional values. Thus, we suggest to estimate the value of gradient on each search point with two ways. One is by using the same gradient estimation method to simulates perturbation method [7]. Another is by using differential values of radial basis function networks to approximate a shape of the objective function.

4.1. gradient estimation by using simulates perturbation method

The estimated value of $\nabla f(\boldsymbol{x}(t))$ is defined as

$$\nabla f(\boldsymbol{x}(t)) \simeq \frac{f(\boldsymbol{x}(t)+\boldsymbol{c}(t)_i) - f(\boldsymbol{x}(t)-\boldsymbol{c}(t)_i)}{2c(t)_i}. \quad (8)$$

Here $c(t)$ is perturbation vector and $c(t)_i$ is its i th element.

In order to show the swarm behavior, we carry out numerical experiment with $\Delta t = 0.1$ and the final time $T = 4000$. The value of gradient is estimated in each step. Fig.5 and 6 show the positions of search points in the final step. The set of parameters is same as in section 3 except ω. Here we also use $\omega = 0.3$. By using estimated values of the gradient, search points gather to local minimum. Thus it seems to be possible to find local minimum. Fig.8–9 are step–position (in a dimension) figures and show the positions of search points. From Fig.8–9, we can find that each search point how go to a local minimum. Moreover fig.10 show the maximum, mean and minimum values of objective function in each step, and fig.11 show the values with $D = 5$. An example of jumping out is shown in the x_1 trajectory of search points in fig.9. It is because a pare of search points come close each other. In this case, the distance between the pare of them so small, then a larger repulsion force from each other is occurred. As a result, a pare of search points jumps out due to such large force. Unstable movement like the above is not acceptable and it is expected to chose the other interaction function to inhibit from such movement. In addition, we find another example that population is divided into two with $\omega = 0.9$ and $r = 1$ but the movements are unstable when the search points are near by local minimums. It seems that the above movements is due to some errors of estimation of gradient. Thus it is necessary to select the value of ω and/or to chose the other way to estimate the gradient.

4.2. gradient estimation by radial basis function network

The estimated value of $\nabla f(\boldsymbol{x}(t))$ is defined as

$$\nabla \left(\sum_k W_k \phi_k(\boldsymbol{x}) \right). \quad (9)$$

Here k is an index of search point, $\phi_k(\boldsymbol{x})$ is kernel function and W_k is network weight. We use Gauss function $\exp(-\frac{1}{2}(\boldsymbol{x}-\boldsymbol{\mu}_k)^T \sum^{-1}(\boldsymbol{x}-\boldsymbol{\mu}_k))$ as kernel function $\phi_k(\boldsymbol{x})$. Here the center of Gauss function $\boldsymbol{\mu}_k$ is assigned to the positions of k th search point. Then we use diagonal matrix which has same variance σ^2_{rbf} in each dimension as the variance–covariance matrix of Gauss function $\sum$. Network weights W_k is given by using LU decompose so as $\sum_k W_k \phi_k(\boldsymbol{x})$ is fit to observed objective functional values.

In order to show the swarm behavior, we carry out numerical experiment with $\Delta t = 0.1$ and the final time $T = 4000$. we use $\omega = 0.9$, $\alpha = 1.0$ and $\gamma = 1.0$ with $\sigma_{rbf} = 2.0, 5.0$. The value of gradient is estimated in 10 steps. Fig.12–14 show final positions of search points, positions in a dimension in each step and objective function values. It seems that search points do not gather near by the local minimums. This behavior may be due to over-trained network weights W_k. So how to give the weights and kernel functions are under consideration.

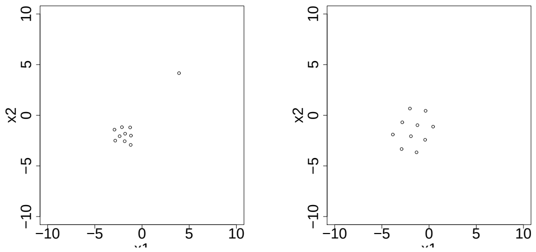

Figure 6: Positions in the final step with $D = 2$, $\omega = 0.3$ and $\alpha = 1$. $r = 1$(left), $r = 2$(right).

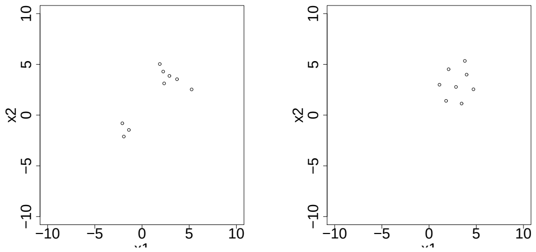

Figure 7: Positions in the final step with $D = 2$, $\omega = 0.3, \alpha = 2$. $r = 1$(left), $r = 2$(right).

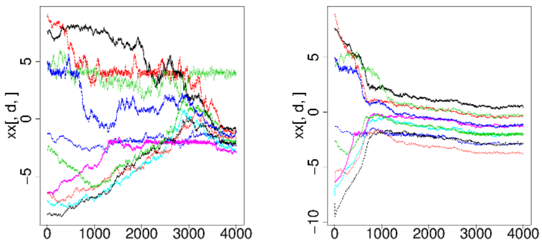

Figure 8: Positions in each step with $D = 2$, $\omega = 0.3$, $\alpha = 1$ and $r = 1$.(left:x_1, right:x_2)

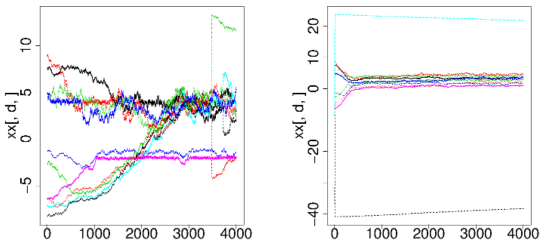

Figure 9: Positions in each step with $D = 2$, $\omega = 0.3$, $\alpha = 1$ and $r = 2$.(left:x_1, right:x_2)

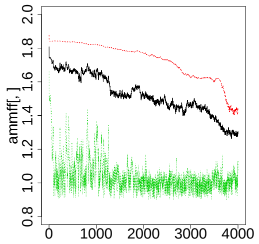

Figure 10: Functional value(maximum, mean, minimum) in each step with $D = 2$, $\omega = 0.3, \alpha = 1$. $r = 1$(left), $r = 2$(right).

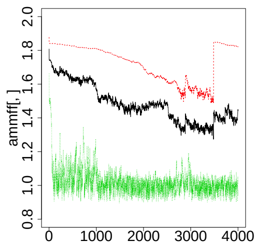

Figure 11: Functional value(maximum, mean, minimum) in each step with $D = 2$, $\omega = 0.3, \alpha = 2$. $r = 1$(left), $r = 2$(right).

5. Conclusion

In this paper, we report the availability of new search method for local minimums based on a mathematical swarming model described by stochastic differential equations. Now only the behaviors with numerical experiments are shown. However we will try to build a new model which include a better interaction force function and a better gradient estimation method.

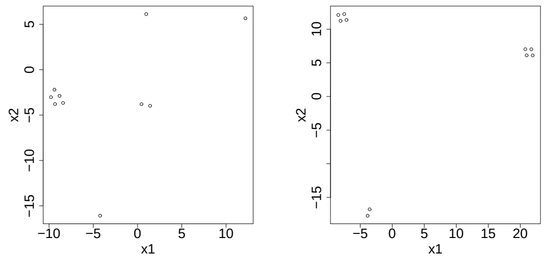

Figure 12: Positions in the final step with $D = 2$, $\omega = 0.9$, $\alpha = 1.0$, $\gamma = 1.0$. $\sigma_{rbf} = 2.0$(left), $\sigma_{rbf} = 5.0$(right).

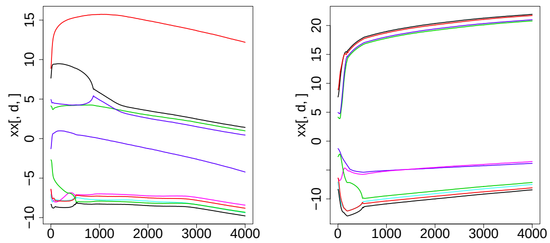

Figure 13: Positions in each step with $D = 2$, $\omega = 0.9$, $\alpha = 1.0$, $\gamma = 1.0$. $\sigma_{rbf} = 2$(left), $\sigma_{rbf} = 5$(right).

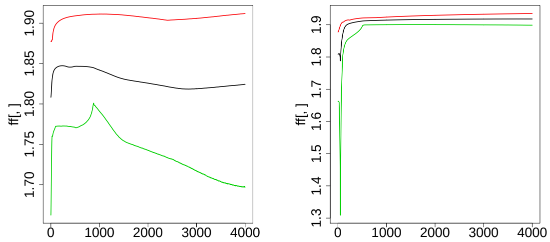

Figure 14: Functional value(maximum, mean, minimum) in each step with $D = 2$, $\omega = 0.9$, $\alpha = 1.0$, $\gamma = 1.0$. $\sigma_{rbf} = 2$(left), $\sigma_{rbf} = 5$(right).

References

[1] M. Dorigo, and L. Gambardella, "Ant Colony Systems: A Cooperative Learning Approach to the Traveling Salesman Problem," *IEEE Transactions on Evolutionary Computation, Vol. 1, No. 1, pp 53–66*, 1997.

[2] Dervis Karaboga, Bahriye Basturk, "A powerful and efficient algorithm for numerical function optimization: artificial bee colony (ABC) algorithm," *Journal of Global Optimization*, Vol. 39, No. 3, pp.459–471, 2007.

[3] Passino, K.M., "Biomimicry of bacterial foraging for distributed optimization and control," *IEEE Control Systems*, Vol.—22, No. 3, pp. 52–67, 2002.

[4] J. Kennedy, and R. Eberhart,"Particle Swarm optimization," *Proceedings of IEEE International Conference on Neural Networks*, pp. 1942–1948, 1995.

[5] Craig W. Reynolds, "Flocks, Herds, and Schools: A Distributed Behavioral Model, in Computer Graphics," *Proceedings of the 14th annual conference on Computer graphics and interactive techniques*, pp. 25–34, 1987.

[6] T. Uchitane, T. V. Ton and A. Yagi, "An ordinary differential equation model for fish schooling," *Scientiae Mathmaticae Japonicae*, 2013.

[7] Y. Maeda, "Simultaneous Perturbation Optimization Methods and Their Applications," *Institute of System, Control and Information Engineers*, Vol. 52. No. 2, pp. 47–53, 2008.(in Japanese)

[8] D. S. Broomhead and D. Lowe, "Radial Basis Functions, Multi-Variable Functional Interpolation and Adaptive Networks," *Royal Signals and Radar Establishment*, No. 4148, 1988.